



## Laboratory evolution reveals a two-dimensional rate-yield tradeoff in microbial metabolism

Cheng, Chuankai; O'Brien, Edward J.; McCloskey, Douglas; Utrilla, Jose; Olson, Connor; LaCroix, Ryan A.; Sandberg, Troy E.; Feist, Adam M.; Palsson, Bernhard O.; King, Zachary A.

*Published in:*  
PLOS Computational Biology

*Link to article, DOI:*  
[10.1371/journal.pcbi.1007066](https://doi.org/10.1371/journal.pcbi.1007066)

*Publication date:*  
2019

*Document Version*  
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

*Citation (APA):*  
Cheng, C., O'Brien, E. J., McCloskey, D., Utrilla, J., Olson, C., LaCroix, R. A., Sandberg, T. E., Feist, A. M., Palsson, B. O., & King, Z. A. (2019). Laboratory evolution reveals a two-dimensional rate-yield tradeoff in microbial metabolism. *PLOS Computational Biology*, 15(6), [e1007066].  
<https://doi.org/10.1371/journal.pcbi.1007066>

---

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

RESEARCH ARTICLE

# Laboratory evolution reveals a two-dimensional rate-yield tradeoff in microbial metabolism

Chuankai Cheng<sup>1</sup>, Edward J. O'Brien<sup>1</sup>, Douglas McCloskey<sup>1</sup>, Jose Utrilla<sup>1</sup>, Connor Olson<sup>1</sup>, Ryan A. LaCroix<sup>1</sup>, Troy E. Sandberg<sup>1</sup>, Adam M. Feist<sup>1,2</sup>, Bernhard O. Palsson<sup>1,2,3</sup>, Zachary A. King<sup>1</sup>\*

**1** Department of Bioengineering, University of California San Diego, La Jolla, California, United States of America, **2** Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Lyngby, Denmark, **3** Department of Pediatrics, University of California San Diego, La Jolla, California, United States of America

\* [zaking@ucsd.edu](mailto:zaking@ucsd.edu)



## Abstract

Growth rate and yield are fundamental features of microbial growth. However, we lack a mechanistic and quantitative understanding of the rate-yield relationship. Studies pairing computational predictions with experiments have shown the importance of maintenance energy and proteome allocation in explaining rate-yield tradeoffs and overflow metabolism. Recently, adaptive evolution experiments of *Escherichia coli* reveal a phenotypic diversity beyond what has been explained using simple models of growth rate versus yield. Here, we identify a two-dimensional rate-yield tradeoff in adapted *E. coli* strains where the dimensions are (A) a tradeoff between growth rate and yield and (B) a tradeoff between substrate (glucose) uptake rate and growth yield. We employ a multi-scale modeling approach, combining a previously reported coarse-grained small-scale proteome allocation model with a fine-grained genome-scale model of metabolism and gene expression (ME-model), to develop a quantitative description of the full rate-yield relationship for *E. coli* K-12 MG1655. The multi-scale analysis resolves the complexity of ME-model which hindered its practical use in proteome complexity analysis, and provides a mechanistic explanation of the two-dimensional tradeoff. Further, the analysis identifies modifications to the P/O ratio and the flux allocation between glycolysis and pentose phosphate pathway (PPP) as potential mechanisms that enable the tradeoff between glucose uptake rate and growth yield. Thus, the rate-yield tradeoffs that govern microbial adaptation to new environments are more complex than previously reported, and they can be understood in mechanistic detail using a multi-scale modeling approach.

## OPEN ACCESS

**Citation:** Cheng C, O'Brien EJ, McCloskey D, Utrilla J, Olson C, LaCroix RA, et al. (2019) Laboratory evolution reveals a two-dimensional rate-yield tradeoff in microbial metabolism. PLoS Comput Biol 15(6): e1007066. <https://doi.org/10.1371/journal.pcbi.1007066>

**Editor:** Vassily Hatzimanikatis, Ecole Polytechnique Fédérale de Lausanne, SWITZERLAND

**Received:** September 17, 2018

**Accepted:** May 2, 2019

**Published:** June 3, 2019

**Copyright:** © 2019 Cheng et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the manuscript and its Supporting Information files.

**Funding:** Funding for this research was provided by the Novo Nordisk Foundation through the Center for Biosustainability at the Technical University of Denmark (NNF10CC1016517). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Author summary

This study reconciles multiple existing microbial rate-yield tradeoff theories with experimental data. There is great interest in developing quantitative descriptions of the

**Competing interests:** The authors have declared that no competing interests exist.

relationship between growth rate and growth yield [1]. However, some reported experiments [2–4] in the literature do not agree with existing theories [5–7]. Specifically, overflow metabolism in *E. coli* can either be coupled [5, 8] or decoupled [2–4] from growth rate. We found that adaptive laboratory evolution (ALE) experiments of *E. coli* reveal a two-dimensional rate-yield tradeoff in adapted strains where the dimensions are (i) a tradeoff between growth rate and growth yield, previously reported by [5], and (ii) a tradeoff between substrate uptake rate and growth yield. The appearance of this two-dimensional tradeoff during adaptation suggests that microorganisms adapting to new environments are subject to a more complex set of rate-yield tradeoffs than previously reported [5, 6]. In this study, the two-dimensional rate-yield tradeoff is quantitatively explained through our multi-scale modeling approach, combining a previously reported small-scale proteome allocation model [5] with a genome-scale model of metabolism and gene-expression (ME-model) [9]. The modeling approach is also instrumental to future studies.

## Introduction

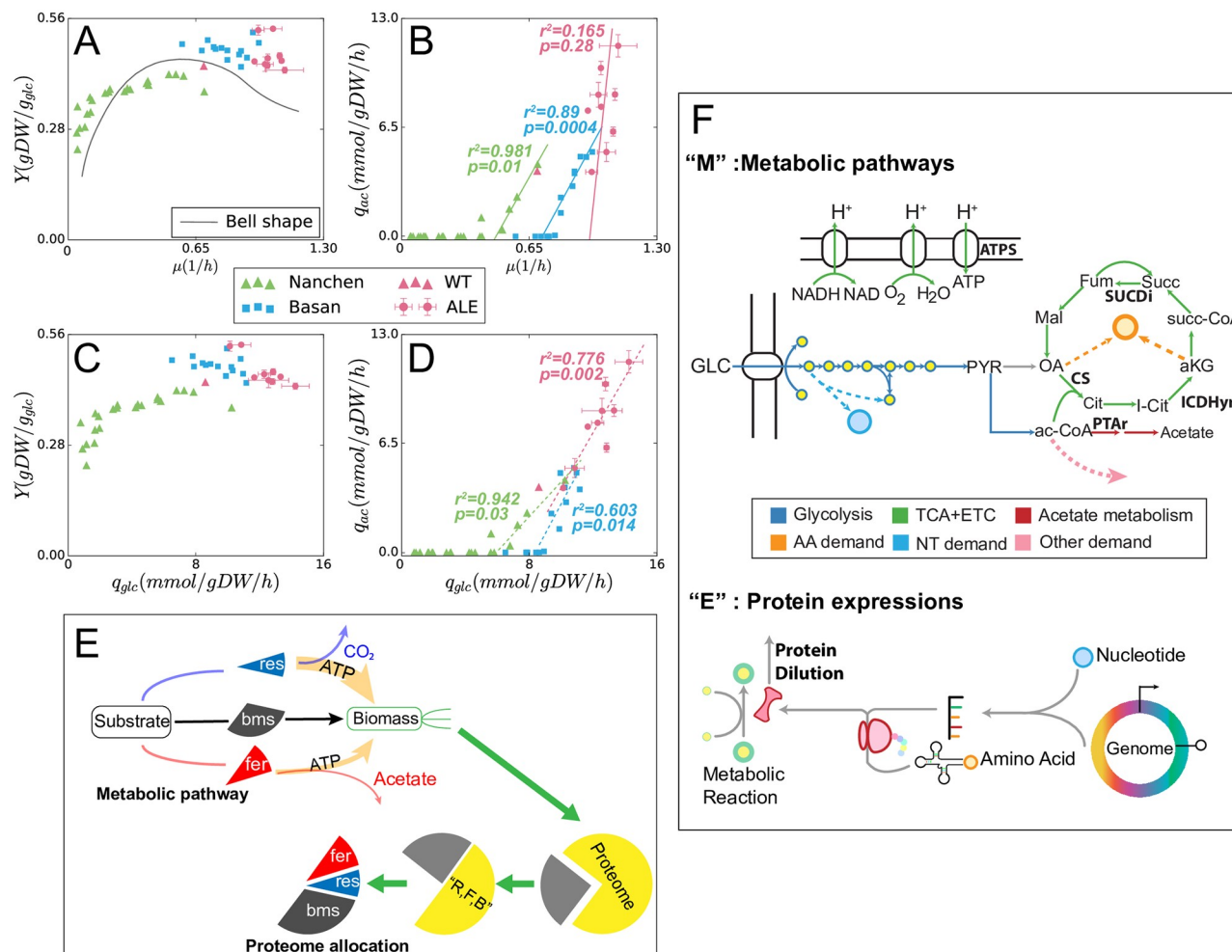
Growth rate and yield are basic features of microbial life that are widely implicated in cell fitness, adaptation, and evolution [1]. The specific growth rate,  $\mu$ , represents the number of doublings of bacterial density per unit time [10]. The yield,  $Y$ , is the ratio between  $\mu$  and the rate of substrate consumption [10, 11]. The mathematical relation between  $\mu$  and  $Y$  can be written as:

$$Y = \frac{\mu}{M_{\text{substrate}} \cdot q_{\text{substrate}}} \quad (1)$$

where  $M_{\text{substrate}}$  is the molecular weight of the substrate and  $q_{\text{substrate}}$  is the substrate uptake rate.

In the context of modeling the phenotypic relation between substrate uptake, metabolism, and biomass growth, there is a great interest in developing quantitative descriptions of the relationship between  $\mu$  and  $Y$ . The wide-ranging measurements of  $\mu$  and  $Y$  (Fig 1A) across microbial communities and environments raised interest into the exact nature of the  $\mu$ – $Y$  relationship [1]. At low  $\mu$ , positive correlations between  $\mu$  and  $Y$  have been observed [8], and these can be explained by non-growth-associated cell maintenance requirements that make slow growth inefficient [11]. At high  $\mu$ , negative correlations between  $\mu$  and  $Y$  are observed [5], and for *E. coli*, this can be explained by a tradeoff between metabolic efficiency and enzymatic efficiency that lead to decreasing  $Y$  at high  $\mu$  [12, 13]. In particular, *E. coli* exhibits a tradeoff between respiration, which has higher energy yield per carbon substrate (more metabolically-efficient), and acetate fermentation, which requires less enzyme per carbon substrate (more proteome-efficient). Therefore, acetate excretion increases linearly with  $\mu$  above a threshold growth rate (green and blue lines in Fig 1B) [5]. [1] summarized these observations where positive  $\mu$ – $Y$  correlation at low  $\mu$  and negative  $\mu$ – $Y$  correlation at high  $\mu$  are different parts of a bell-shaped  $\mu$ – $Y$  curve (Fig 1A). However, recent experiments suggest that adaptation to new environments can modify the bell-shaped  $\mu$ – $Y$  tradeoff [3, 14, 15].

Microorganisms rapidly adapting to environmental niches [17, 18], and adaptation mechanisms can be studied directly through adaptive laboratory evolution (ALE) [19]. When strains are adapted through ALE for growth in a liquid minimal medium, they achieve higher  $\mu$  compared to the wild-type (Fig 1A), ALE-adapted strains have been shown to rapidly acquire regulatory mutations that modify proteome allocation, but they do not acquire new metabolic



**Fig 1. *E. coli* growth phenotypes in minimal media and multi-scale modeling approaches.** Data on the plots are recorded in S7–S10 Tables (Supporting Information). (A–D) Phenotypic data for *E. coli* strains including  $Y$ ,  $\mu$ ,  $q_{ac}$  and  $q_{glc}$  data. The  $Y$  is calculated by  $\frac{\mu}{M_{glc} \cdot q_{glc}}$ , where the molecular weight of glucose  $M_{glc} = 180.156 \text{ g/mol}$ . Two datasets are presented from literature, for chemostat growth [8] (green triangles) and substrate titration [5] (blue squares). These are compared to strains adapted for maximum growth rate through ALE (this study; red circles; error bars for standard deviation across duplicates). The bell-shaped  $\mu$ – $Y$  relationship proposed by [1] is included for reference. (E) Diagram of the SSME-model derived from [5]. The model consists of three pathways: respiration (R, res) and fermentation (F, fer) generate different amounts of energy, feeding the biomass (B, bms) pathway to synthesize biomass. (F) Diagram of the genome-scale ME-model that includes a genome-scale reconstruction of metabolic pathways (“M”) and protein expression machinery (“E”) [9, 16]. Only central carbon metabolism is sketched, other biosynthesis pathways (AA amino acids, NT nucleotides and others) are simplified as dashed arrows in the plot, but remain fine-grained in the genome-scale ME-model that we analyze in this study. For the “E” pathway, it starts with the amino acids and nucleotides that are synthesized through “M” pathway, and the end product of “E” pathway is the dilution of the protein (enzyme). The dilution rate of the enzyme determines the corresponding metabolic reaction rate with the factor  $k_{eff}$ . More details about  $k_{eff}$  constraints can be found in “3 Proteome constraints in the ME-model” in S1 Appendix.

<https://doi.org/10.1371/journal.pcbi.1007066.g001>

capabilities within the time frame of reported short-term (4 to 8 weeks) adaptation experiments [3, 15, 20]. By analyzing ALE-adapted strains, we can reveal the strategies that allow cells to optimize their proteome allocation for growth in an environmental niche, subject to the constraints of their metabolic capabilities (i.e. their repertoire of pathways) and constraints on the kinetic efficiencies of their enzymes [3, 20, 21].

Contrary to the negative  $\mu$ – $Y$  relationship at high  $\mu$  observed in wildtype strains, the endpoint strains of ALE experiments of *E. coli* selected for high  $\mu$  in a minimal medium do not have  $\mu$ – $Y$  data points aligning on the bell-shape curve in Fig 1A, but reveal an uncorrelated

relationship between  $\mu$  and  $Y$  [3, 15]. In these experiments, strains exhibited little variation in  $\mu$  but high variation in  $Y$  and acetate excretion rate  $q_{ac}$ . Previous studies similarly reported that overflow metabolism can be nearly eliminated through genetic engineering without any effect on growth rate in *E. coli* [2, 4]. Thus, the negative  $\mu$ - $Y$  correlation at high growth rates does not appear to be a fundamental constraint on fast-growing cells. A mechanistic model of the full  $\mu$ - $Y$  relationship must be able to reconcile the bell-shaped curve observed for individual strains with the uncorrelated  $\mu$ - $Y$  phenotypes seen in ALE-adapted strains (Fig 1A).

A number of theoretical and computational models have been developed to describe rate-yield tradeoffs. For the positive  $\mu$ - $Y$  correlation, maintenance requirements can be quantitatively described using algebraic growth laws [8, 11]. The maintenance requirement has been modeled as non-growth associated maintenance (NGAM) in the genome-scale models (GEMs) of metabolism, which can be simulated as an optimization problem, predicting  $\mu$  and  $Y$  when substrate uptake rates (e.g.  $q_{glc}$ ) are known [22]. For the negative  $\mu$ - $Y$  correlation, quantitative models of overflow metabolism have been developed [5–7]. In particular, quantitative measurements of *E. coli* growth in well-controlled environments revealed a linear-threshold response of acetate excretion ( $q_{ac}$ ) with increasing  $\mu$  [5]. To represent the full range of the  $\mu$ - $Y$  relationship, a constraint allocation flux balance analysis model (CAFBA) was reported that combines a GEM with proteome allocation constraints [6]. A similar solution can be formulated from a bottom-up reconstruction of metabolism and macromolecular expression (ME-model, [9]) that incorporates the protein synthesis pathways into a GEM and applies coupling constraints related to enzyme kinetics parameters on each individual reaction. However, none of these models have been used to explain experiments where  $\mu$  and  $Y$  are decoupled through laboratory evolution or genetic engineering.

In this study, we show that the wide range of  $\mu$ - $Y$  observations in *E. coli* can be explained by a two-dimensional rate-yield tradeoff, where the first dimension is the characteristic  $\mu$ - $Y$  tradeoff associated with acetate overflow metabolism and the second dimension is a tradeoff between glucose uptake rate ( $q_{glc}$ ) and  $Y$  that appears during ALE adaptation. We employ a multi-scale modeling approach to provide a mechanistic description of the two-dimensional rate-yield tradeoff. By deriving the relationship between the ME-model and the previously reported small-scale proteome allocation model [5], we are able to develop a workflow for modifying ME-model parameters to fit experimental data, and we achieve quantitative predictions for simulations of  $\mu$ - $Y$  (the first dimension of the rate-yield tradeoff). This multi-scale modeling approach predicts a two-dimensional rate-yield tradeoff, and it suggests that the second dimension of the tradeoff can be explained by changes in P/O ratio and the flux balance between glycolysis and pentose phosphate pathway. This multi-scale modeling approach predicts the systemic response of the cell to growth selection, representing the relationships between P/O ratio, glycolytic-PPP flux balance, and the two dimensions of the rate-yield tradeoff.

## Results

### Adaptive laboratory evolution reveals a two-dimensional rate-yield tradeoff

To explore the metabolic constraints on *E. coli* growth, adaptive laboratory evolution (ALE) was used to adapt *E. coli* K-12 MG1655 to maximize growth at 37°C in a liquid culture with a minimal medium containing glucose [3]. Eight independent experiments were performed on an automated ALE platform to achieve  $8.3 \times 10^{12}$  to  $18.3 \times 10^{12}$  cumulative cell divisions [23]. Phenotype characterization was performed on eight ALE endpoint strains, including

quantitative measurements of  $\mu$ ,  $q_{glc}$ ,  $q_{ac}$ , and other common metabolic byproducts of *E. coli* ([Materials and methods](#)).

A diversity of metabolic phenotypes was observed in the ALE endpoint strains. Through ALE,  $\mu$  increased from  $0.7 \text{ h}^{-1}$  for wild-type (red triangles in [Fig 1A–1D](#)) to  $0.95\text{--}1.10 \text{ h}^{-1}$  (red circles with error bars in [Fig 1A–1D](#)). Based on previous reports, we expected a linear relationship between  $\mu$  and  $q_{ac}$ . However, ALE endpoint strains achieved a wide ranging  $q_{ac}$  from  $3.9\text{--}11.4 \text{ mmol gDW}^{-1} \text{ h}^{-1}$  (where wild-type  $q_{ac}$  was  $3.9 \text{ mmol gDW}^{-1} \text{ h}^{-1}$ ). While we did not observe a correlation between  $\mu$  and  $q_{ac}$  in these strains ([Fig 1B](#)), there was a clear correlation between  $q_{glc}$  and  $q_{ac}$  ([Fig 1D](#)).

Two of the ALE endpoint strains with similar  $\mu$  (3% difference) but distinct  $Y$  (30% difference) have been processed for  $^{13}\text{C}$  metabolic flux measurements ([S11 Table](#)). The measured metabolic fluxes using  $^{13}\text{C}$  metabolic flux analysis ( $^{13}\text{C}$  MFA, see [“Materials and methods”](#)) shows the positive correlation between TCA fluxes,  $q_{TCA}$ , and  $Y$  for the ALE strains. For the ALE strain with larger  $Y$ ,  $q_{glc}$  and  $q_{ac}$  are lower and  $q_{TCA}$  is higher. Therefore, for a fixed  $\mu$ ,  $Y$  increases as  $q_{TCA}$  increases and  $q_{ac}$  decreases, indicating a pathway switch between the TCA cycle and acetate overflow depending on  $q_{glc}$ .

Therefore, combining with the referenced study [5], for a wild-type strain, there is a  $\mu$ – $Y$  tradeoff. And for the isogenic ALE strains, a  $q_{glc}$ – $Y$  tradeoff appears. For both tradeoffs,  $Y$  varies with the pathway switch between TCA cycle and acetate overflow. In this paper, we call the  $\mu$ – $Y$  and  $q_{glc}$ – $Y$  tradeoffs a two-dimension rate-yield tradeoff, since they are tradeoffs between different “rates” (growth rate  $\mu$  and glucose uptake rate  $q_{glc}$ ) and the same yield (glucose yield  $Y$ ), and they share the same phenotypic behavior of TCA–acetate overflow pathway switch.

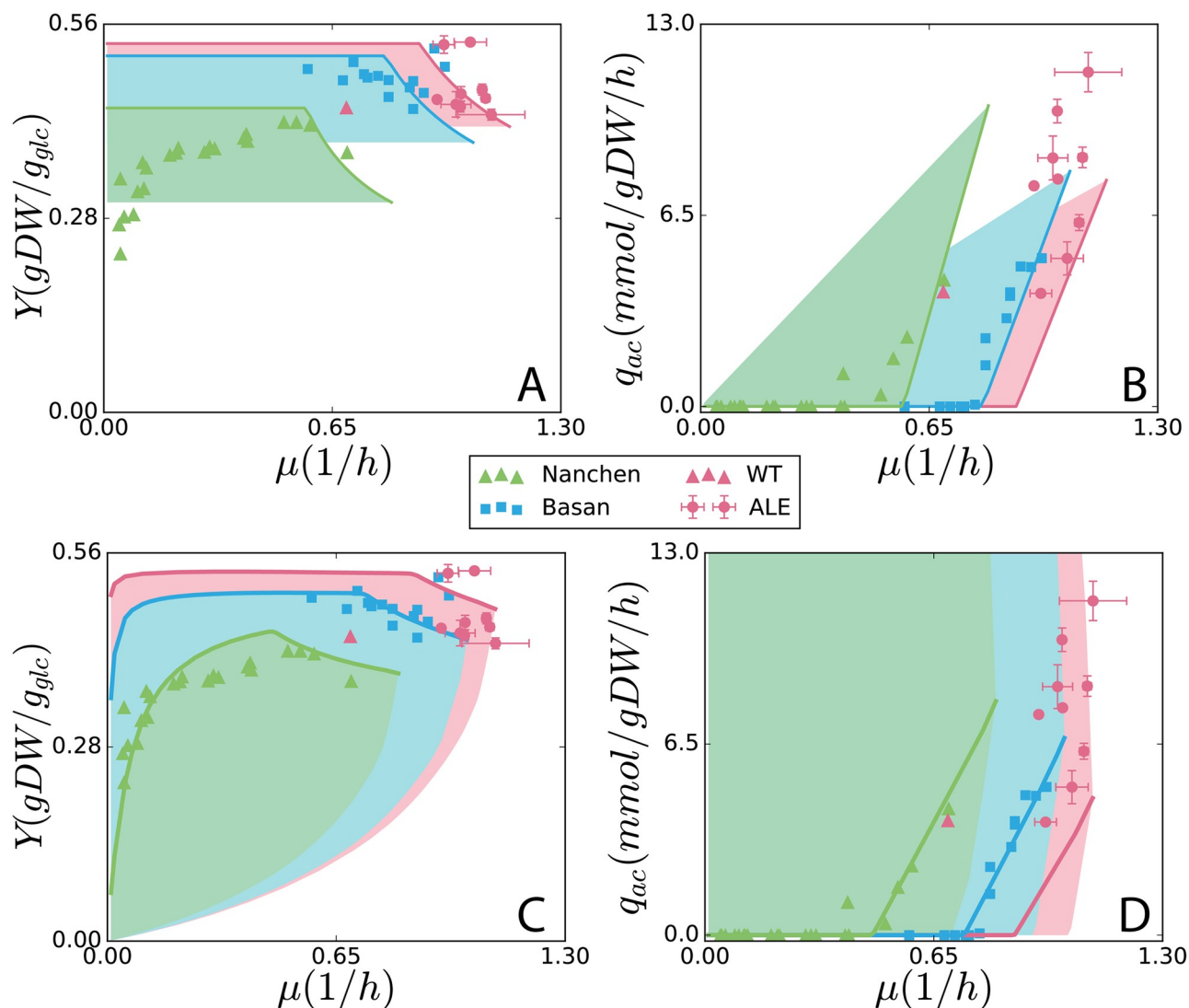
Correlations between  $q_{glc}$ ,  $Y$ , and  $q_{ac}$  have been observed previously for *E. coli* strains [2–4], and moreover, a bacterial engineering approach has been reported to vary  $q_{ac}$  by manipulating the substrate uptake system [24]. In one of these studies, [4] showed that switching electron transport chain (ETC) enzyme selection (and thereby modifying the P/O ratio) can cause a  $q_{glc}$ – $Y$  tradeoff at a low  $\mu$  of  $0.15 \text{ h}^{-1}$ . ALE gained  $q_{ac}$  and  $Y$  decoupled from  $\mu$ , which seemingly differs from the reported correlation between  $\mu$ – $Y$  and  $\mu$ – $q_{ac}$  [5, 8]. The ME-model used in this study simulates the relationships between these  $q_{glc}$ – $q_{ac}$  and  $q_{glc}$ – $Y$  tradeoffs, connecting to the mechanisms of  $\mu$ – $Y$  tradeoffs (the bell-shaped curve in [Fig 1A](#)) by established models [5, 6].

To enable our analysis, it is important to note that ALE endpoint strains rapidly acquire regulatory mutations, but they do not acquire new metabolic capabilities within the time frame of these experiments [3, 15, 20]. The linear correlation between  $q_{ac}$  and  $\mu$  reported previously was identified for an isogenic strain [5, 8]. In contrast, our observations of a decoupling between  $q_{ac}$  and  $\mu$  appear when comparing adapted strains. However, because these adapted strains have only regulatory mutations, their phenotypes represent the limits of what *E. coli* cells can achieve while bounded by metabolic and proteomic constraints (but not by regulation). This type of adaptation and the associated phenotypic tradeoffs are useful for understanding cellular adaptation to ecological niches where regulatory adaptation can occur rapidly [18].

## ME-model data fitting with a multi-scale modeling approach

To explain these experimental observations, we sought a modeling approach that could quantitatively predict the  $\mu$ – $Y$  and  $\mu$ – $q_{ac}$  relationships. Our modeling approach starts with fitting the linear-threshold (blue line in [Fig 1B](#))  $\mu$ – $q_{ac}$  relation [5] using the framework of ME-model [9, 16]. We first considered a previously reported coarse-grained model of proteome allocation [5] that describes *E. coli* overflow metabolism ([Fig 1E](#)). [5] solves  $Y$  and  $q_{ac}$  as functions of  $\mu$ , and assuming that the cells pick the maximum  $Y$  under each particular  $\mu$ . This indicates that





**Fig 2. SSME-model and ME-model simulations.** Growth phenotypes of *E. coli* from simulations: (A, B) using the SSME-model and (C, D) ME-model. Simulations were fit to experimental data for each of the three datasets, K-12 MG1655 chemostat [8] (green triangles), NCM3722 substrate titration [5] (blue squares), and strains adapted from wild-type K-12 MG1655 (red triangle, [3] for maximum growth rate through ALE (this study, red circles, error bars for standard deviation across duplicates). The Y-maximized solutions are displayed as solid lines in all plots. Solution spaces are simulated by taking the feasible range between maximum ( $Y_{max}$ - $Y_{min}$  in A and C,  $q_{ac,max}$ - $q_{ac,min}$  in B and D) and minimum values. For both models, fitting was performed by manipulating three global parameters: unmodeled protein fraction (UPF), growth-associated maintenance (GAM), and non-growth associated maintenance (NGAM). Details of the fitting approach are provided in Materials and Methods.

<https://doi.org/10.1371/journal.pcbi.1007066.g002>

high- $Y$  growth strategies have a fitness benefit in spatially structured environments, for instance, the wild-type cultures collected from colonies, that has been demonstrated through a  $Y$ -selection system [14], and more efficient strategies also leave more resources for cells that are hedging against future stresses [20]. The evolutionary history of *E. coli* includes growth in structured environments and a wide range of stresses that could have placed a selection pressure on increasing  $Y$ . Therefore, we focused on fitting the observed wild-type chemostat [8] and uptake titration [5] data for the  $Y$ -maximized growth solution (green and blue data points in Fig 2).

The coarse-grained proteome allocation model [5] was intended to make predictions at high  $\mu$  and thus only captures the negative  $\mu$ - $Y$  relation (Fig 2A). The parameters in the coarse-grained model have a strong experimental basis in fine-grained protein abundances measurements in high growths, and the model simulates accurate predictions of  $\mu$ - $q_{ac}$  [5].

We also considered the genome-scale ME-model *iJL678*-ME [16]. With the default parameter settings in the ME-model, simulations had a poor quantitative prediction [9] of  $\mu$ - $q_{ac}$  to the uptake titration data (S3F Fig). As [5] has shown experimentally that the overflow metabolism is fundamentally caused by the tradeoff between metabolic efficiency (reaction stoichiometry) and protein efficiency (enzyme turnover rate). Since the reaction stoichiometry in the ME-model has been mass-balanced and well established, we suspect that this poor fit from the ME-model can be explained by inaccurate genome-wide enzyme turnover rates ( $k_{eff}$ s) that ME-model researchers have been seeking to improve [16, 25, 26]. We sought to modify the  $k_{eff}$ s to fit the  $\mu$ - $q_{ac}$  data. However, since each of the 5266 reactions in the genome-scale ME-model has a  $k_{eff}$  parameter, it is difficult to directly fit the parameters to measured data.

Therefore, we pursued a multi-scale modeling approach where the coarse-grained model was used to analyze the effects of proteome-efficiency at the level of coarse-grained pathways instead of each individual reaction, which helps to tune the fine-grained parameters in the ME-model. To connect the coarse-grained and fine-grained models, we first found that the proteome efficiency ( $\epsilon$ ) parameters in the coarse-grained model share a conceptual basis with the enzyme efficiency parameter  $k_{eff}$ s in ME-models (“3 Proteome constraints in the ME-model” in S1 Appendix). Thus, we were able to reformulate the coarse-grained model within the framework as the ME-model (S1 Fig). The resulting small-scale ME-model (SSME-model) has parameters directly analogous to those in the genome-scale ME-model (See “5 SSME-model parameters derivation” and “6 Matlab and COBRAme implementation” in S1 Appendix). The resulting SSME-model generates identical  $\mu$ - $Y$  and  $\mu$ - $q_{ac}$  predictions to the proteome allocation model.

The SSME-model is a good tool for  $k_{eff}$  parameter sensitivity analysis [27], which provides insights on how to modify the  $k_{eff}$  of the ME-model to achieve quantitative fit. As a result, we gained predictions for  $\mu$ - $Y$  and  $\mu$ - $q_{ac}$  from both the SSME- and ME-models (blue curves in Fig 2). Details of the ME-model modifications are in the S1 Appendix (“8 Experimental data fitting”). In summary, with the multi-scale modeling approach, we identified the reactions whose enzymes turnover rates are too high to match the observed phenotypes. Those reactions are involved in different pathways, including the TCA cycle, Entner-Doudoroff pathway, glyoxylate shunt, nucleotide salvage, and fatty acids metabolism (S3 Table). Three global parameters, unmodeled protein fraction (UPF), growth-associated maintenance (GAM), and non-growth-associated maintenance (NGAM) (S2 Table) were then used to predict the phenotype from different strains (green, blue and red curves in Fig 2).

The reason for only modifying global parameters to simulate the ALE adaptation is that the mutations in the ALE strains do not directly related to the enzyme turnover rate ( $k_{eff}$  value) of a particular metabolic reaction. According to previous ALE studies [3], most mutations occur in genes associated with regulations or translations. Even in the cases where mutations might directly change a  $k_{eff}$ , this is hard to model. Therefore, rather than exploring the mechanistic effects of ALE mutations, we focused on the phenotypic changes in the endpoint strains. Some recent studies have shown how individual mutations can have wide-reaching effects on gene expression, metabolic pathway activity, and cell phenotype [3, 20].

The most obvious difference between the SSME-model derived from [5] and ME-model for these phenotypic predictions is the expanded solution space of the ME-model (Fig 2). However, much of the ME-model solution space corresponds to very low yield metabolic solutions. If  $Y$  is maximized during simulations of the SSME-model and ME-model (achieved



by minimizing  $q_{glc}$  at a given  $\mu$ ), the resulting predictions are more similar between the models and lie closer to experimental data (solid blue curves in Fig 2). For the growth-rate-dependent  $Y$ -maximized solutions (solid curves in all 4 panels), though the  $q_{ac}$  lines look similar, the  $Y$  lines look very different in the low growth regime. Where the SSME-model predicts a constant high  $Y$ , the ME-model predicts an initially low  $Y$  that increases rapidly with  $\mu$ . This is because the additional non-growth maintenance energy (NGAM) added in the ME-model. We can also see that in Fig 2C, among the three different solution curves from the ME-model, the curve with lower NGAM has a higher  $Y$  at low  $\mu$ . The NGAM parameters for the different curves are shown in S1 Table in Supporting information. In fact, by adding complexity to the SSME-model or simplifying the ME-model, many intermediate models can be built.

## The ME-model predicts phenotypic diversity in ALE strains

As a result of data fitting, we achieved a quantitative fit of chemostat [8] and batch [5] uptake titration data with the  $Y$ -maximized ME-model solutions (blue and green curves in Fig 2C and 2D). The ALE-adapted strains (red circles in Fig 2) do not align well with the  $Y$ -maximizing solutions (red curves in Fig 2), but they are encompassed by the ME-model solution space. Further analysis of these ALE data points and the corresponding ME-model solutions were used to understand the phenotypic diversity of these adapted strains.

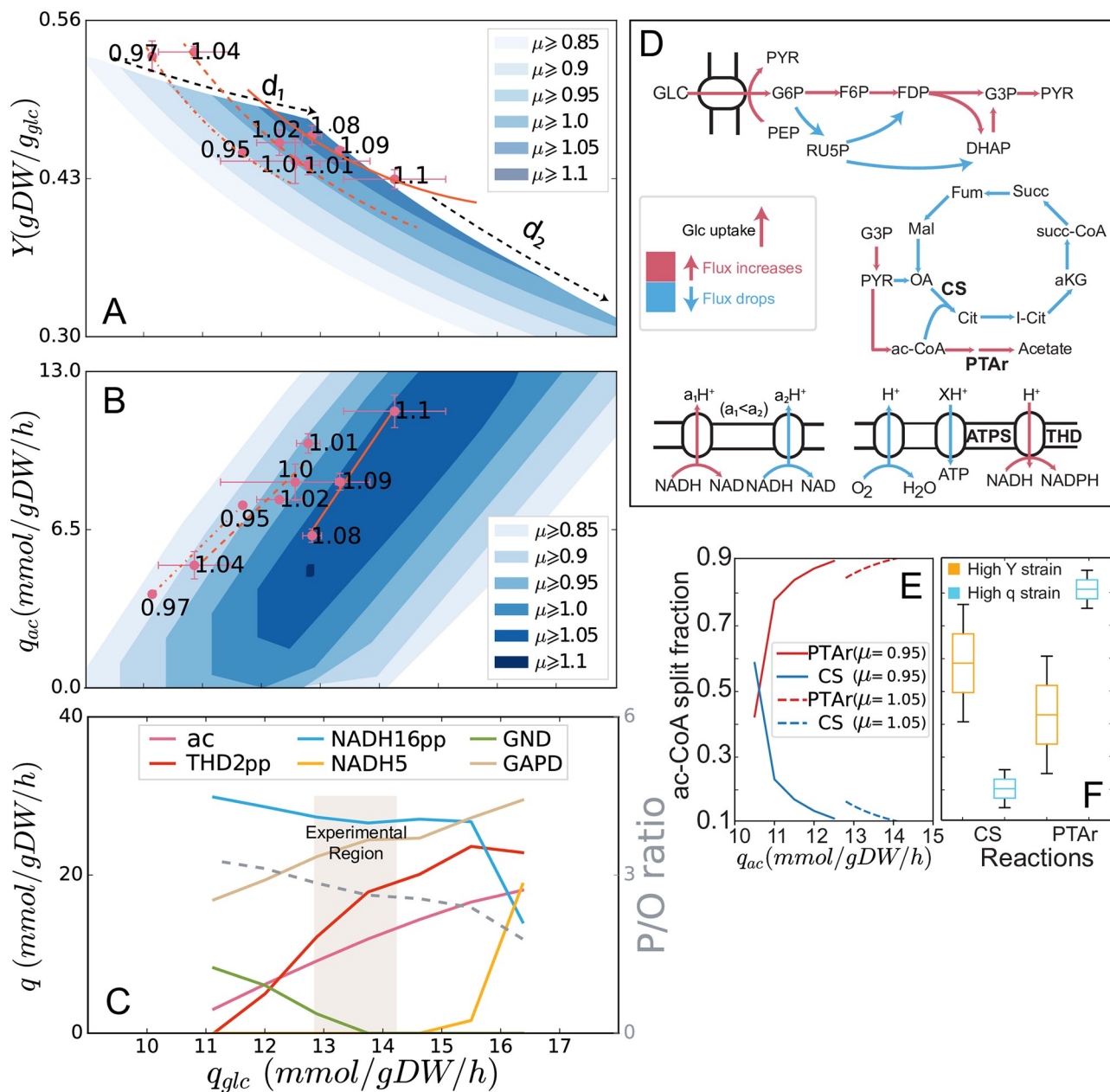
Feasible solutions other than the  $Y$ -maximized solution are achieved through the activation of alternative metabolic pathways which are sub-optimal. The SSME-model does not capture the ALE data points with high  $q_{ac}$  (red region in Fig 2B), while the genome-scale ME-model does (red region in Fig 2D). Moreover, the ME-model predicts feasible growth at lower  $Y$  in the  $\mu$ - $Y$  solution space than the SSME-model. We sought to determine which pathways are responsible for the lower  $Y$  and higher  $q_{ac}$  in ME-model that was not captured by the SSME-model.

Removing reactions from the ME-model can decrease the size of the solution space (S5, S6 and S7 Figs, “8 Solution space variation” in S1 Appendix), making the solution space more similar to the SSME-model solution space. We employed a workflow to identify 24 reactions (S6 Table) that are not activated in the  $Y$ -maximized solutions but are used to enable higher  $q_{ac}$  at lower  $Y$ . We observed that these 24 reactions are part of metabolically inefficient pathways that are alternatives to the  $Y$ -optimal pathways. By extension, metabolically inefficient pathways can be added to the SSME-model to increase the size of the solution space (S7 Fig), making it more similar to the ME-model solution space. Thus, the modified SSME-model can achieve low  $Y$  (S7A Fig) at high  $q_{ac}$  (S7C Fig). Therefore, the difference in predictions of ME-model from the SSME-model is a result of the greater range of metabolic capabilities of the genome-scale model.

## The two-dimensional rate-yield tradeoff

We can now put forward a theory to connect the correlations in  $\mu$ - $Y$  (Fig 1A) (and the associated acetate curve in  $q_{ac}$ - $Y$ , Fig 1B) with the negative correlation in  $q_{glc}$ - $Y$  (Fig 1C) and positive  $q_{glc}$ - $q_{ac}$  correlations (Fig 1D).

To see the relationship between the three variables  $\mu$ ,  $q_{glc}$ , and  $Y$  we generated ME-model solution spaces in  $q_{glc}$  and  $Y$  at increasing lower bounds of  $\mu$  (Fig 3A). These solution spaces represent the flexibility in the model to achieve a particular growth rate. At the  $Y$ -maximized limit of these solution space, we see the established negative  $\mu$ - $Y$  tradeoff where increasing growth rate requires increasing  $q_{glc}$  and decreasing  $Y$  (dashed arrow marked as “d1” in Fig 3A) coupling with increasing  $q_{ac}$  (top edges of solution spaces in Fig 3B). This is the first dimension of the rate-yield tradeoff, “d1”.



**Fig 3. Analysis of the second dimension of the rate-yield tradeoff.** In (A) and (B), the growth rate of each ALE endpoint strain is labeled as black number beside the corresponding data point. (A) Two dimensions of the rate-yield tradeoff. The first dimension “d1” is the negative  $\mu$ -Y correlation at maximum Y, and the second dimension “d2” is the negative  $q_{glc}$ -Y correlation at a fixed  $\mu$ . These correlations are observed in ME-model simulations and experimental data from ALE strains. (B) A correlation between  $q_{glc}$  and  $q_{ac}$  is also observed at fixed  $\mu$  in both the ME-model and ALE endpoint data. Linear fits for the experimental data at similar growth rates are shown as dash-dotted ( $\mu = 0.95$ – $0.97$  h<sup>-1</sup>), dashed ( $\mu = 1.00$ – $1.04$  h<sup>-1</sup>), and solid ( $\mu = 1.08$ – $1.10$  h<sup>-1</sup>) orange lines. These fits are described by the upper edges of the  $q_{glc}$ - $q_{ac}$  solution space at fixed  $\mu$ . For growth between  $1.00$  and  $1.04$  h<sup>-1</sup>,  $r^2 = 0.931$  and  $p = 0.035$ . For growth between  $1.08$  and  $1.1$  h<sup>-1</sup>,  $r^2 = 0.986$  and  $p = 0.071$ . (C) The reaction fluxes in ME-model simulations along the upper edge (maximizing  $q_{ac}$ ) of the solution space for  $\mu = 1.05$  h<sup>-1</sup> in (B). Notably, the P/O ratio (ratio of ATPS flux and oxygen uptake flux, gray dashed curve) is decreasing with increasing  $q_{glc}$ . (D) Variation of fluxes distribution under the same  $\mu$  in central metabolism and electron transport chain. (E) Model simulation: the ac-CoA split fraction to TCA cycle (CS) and acetate fermentation (PTAr). Here, two growth rates ( $0.95$  h<sup>-1</sup> and  $1.05$  h<sup>-1</sup>) are picked for simulation. (F) Experimental verification of the ac-CoA split through <sup>13</sup>C-MFA data, details of the <sup>13</sup>C-MFA data are illustrated in “<sup>2</sup> <sup>13</sup>C metabolic flux analysis” in S1 Appendix. **Abbreviations:** CS: Citrate synthase (*gltA*); PTAr: Phosphotransacetylase (*pta* and *eutD*); ac: acetate excretion; THD2pp: NAD(P) transhydrogenase (catalyzed by the gene product of *pntAB*); NADH16pp: NADH dehydrogenase (*nuoA-N*); NADH5: NADH dehydrogenase (*ndh*); GND: Phosphogluconate dehydrogenase (*gnd*); GAPD: Glyceraldehyde-3-phosphate dehydrogenase (*gapA*); ATPS: ATP synthase (*atpA-I*).

<https://doi.org/10.1371/journal.pcbi.1007066.g003>

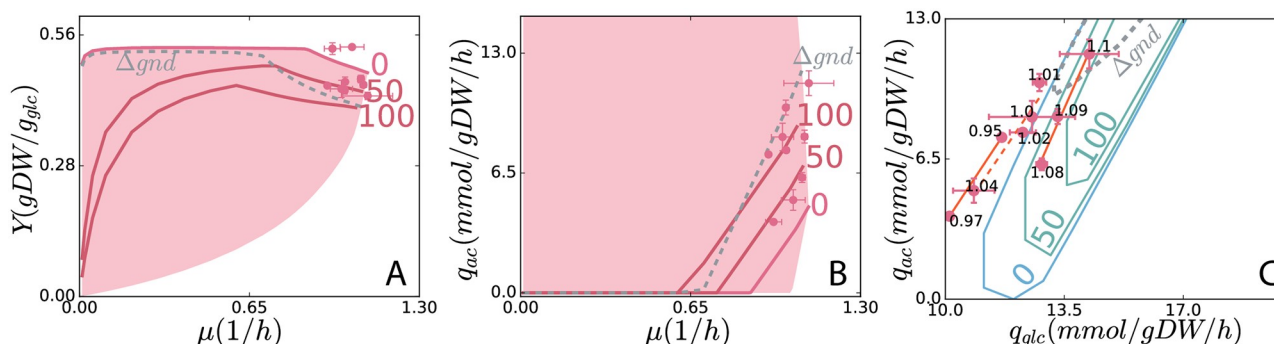
Considering only “d1”, one would expect the acetate production rate in all strains to be fully defined by the growth rate. In the case of this 1-dimensional tradeoff, all points in Fig 3A would appear on the line at the top of the blue solution spaces (parallel to the dotted “d1” line). However, we observed another degree of freedom in the phenotypic space. At a given  $\mu$ , evolved strains can acquire higher  $q_{glc}$ , higher  $q_{ac}$ , and lower  $Y$ . The “d2” tradeoff is defined by a linear correlation in  $q_{glc}$ – $q_{ac}$  (Fig 3B) and a corresponding inverse proportional tradeoff in  $q_{glc}$ – $Y$  (Fig 3A).

The “d2” tradeoff is also predicted by ME-model simulations. At a given  $\mu$ , the ME-model solution spaces extend toward lower  $Y$  and higher  $q_{glc}$ , revealing this inverse proportional relationship in  $q_{glc}$ – $Y$ . The second dimension “d2” can also be seen in  $q_{ac}$ – $q_{glc}$  where the ME-model predicts the  $q_{ac}$ – $q_{glc}$  correlation observed in ALE endpoint as the  $q_{ac}$ –maximized edges of the solution spaces (Fig 3B). The solution spaces predicted by ME-model show broad feasible ranges of acetate production  $q_{ac}$  at a given  $q_{glc}$  and  $\mu$  (“bold” solution spaces in Fig 3B), so the  $q_{glc}$ – $Y$  tradeoff is not required by the model. On the other hand, the relationship between  $q_{glc}$  and  $Y$  is a strict tradeoff in the model (“thin” solution spaces in Fig 3A). Therefore, the ME-model suggests that  $q_{glc}$ – $Y$  is the more fundamental second dimension of the rate-yield tradeoff. To verify that hypothesis, one would look for mutant strains where  $q_{glc}$  increased while the other three phenotypic variables remained fixed (a shift to the right in Fig 3B).

## Mechanisms for the additional rate-yield tradeoff

We sought to identify the particular alternate metabolic strategies in the ME-model that could enable a  $q_{glc}$ – $Y$  tradeoff by identifying the differential pathway usage at a fixed high  $\mu$  (1.05 h<sup>-1</sup> in the ME-model (Fig 3C). The model predicts that when  $q_{glc}$  increases from the  $Y$ -maximized state (minimum  $q_{glc}$ ), flux through the proton-coupled NAD(P) transhydrogenase increases (reaction THD2pp, catalyzed by *pntAB*). In addition, a pathway switch between two different NADH dehydrogenase reactions, NADH5 (*ndh* and NADH16pp (*nuo*, appears at high  $q_{glc}$ . In fact, each of or any combination of the 24 reactions in S6 Table can be activated in the ME-model to achieve high  $q_{glc}$ , high  $q_{ac}$ , and low  $Y$ . There are two common threads among these pathway activations. First, they all reduce the P/O ratio in the simulations (Fig 3C). NADH5 transports fewer protons to the periplasm per electron than NADH16pp. And increasing THD2pp flux drains the proton gradient without contributing to ATP production, thereby reducing P/O ratio (Fig 3D). Second, with the activation of those 24 reactions, glycolytic flux increases (Fig 3D) and pentose phosphate pathway flux decreases (Fig 3C). By comparing to the <sup>13</sup>C metabolic flux analysis (Fig 3E and 3F), the ME-model shows quantitative predictive power for the second-order rate-yield tradeoff.

Experiments that introduce proton leakage have shown a shift towards high  $q_{ac}$  and low  $Y$  [5] in the same  $\mu$ . It has also been shown that the variation of P/O ratio can uncouple the regulation of cytochrome oxidase from the cellular ATP demand [4]. More broadly, energy dissipation through proton leakage is known to be a method of metabolic control in bacteria [28, 29]. To clarify the effect of decreasing of P/O ratio in the ME-model, we added a reaction in the model representing proton leakage (Methods). As a result, we see the  $Y$ -maximized solution with decreased P/O ratios have higher  $q_{glc}$ , higher  $q_{ac}$ , and lower  $Y$  at a given  $\mu$  (Fig 4). Finally, experiments have shown that knocking out *gnd* leads to increased  $q_{glc}$  and  $q_{ac}$  and decreased  $Y$  with little change in  $\mu$  [30]. The ME-model also predicts that *gnd* knockout mutants (“*gnd* knockout simulation” Methods) will have increased  $q_{glc}$ ,  $q_{ac}$  and decreased  $Y$  (Fig 4). Since the ALE experiments do not introduce leaky proton or knock out any genes, it is also possible that multiple mechanisms working together, where the ME-model points to the systemic



**Fig 4. The second-order rate-yield tradeoff demonstrated by decreasing the P/O ratio and knocking out *gnd* (“ $\Delta gnd$ ”) in ME-model simulations.** The drop of P/O ratio is achieved by inducing the proton leakage (“PL”, a pseudo reaction of proton leakage added in the ME-model, details in “Materials and methods.”) reaction flux, as 0, 50 mmol gDW<sup>-1</sup> h<sup>-1</sup> (labeled “50”), and 100 mmol gDW<sup>-1</sup> h<sup>-1</sup> (labeled “100”). The new Y-maximized solution curves (solid red for “PL” flux variation, dashed grey for  $\Delta gnd$ ) in the (A)  $\mu$ -Y and (B)  $\mu$ - $q_{ac}$  solution spaces. (C) The  $q_{ac}$ - $q_{glc}$  solution space contours (fixed  $\mu = 1.0$  h<sup>-1</sup>, solid blue for “PL” flux variation, dashed grey for  $\Delta gnd$ ) were simulated in the ME-model. Growth rates of the experimental strains are labeled as black numbers right next to each data point. The data points in similarly closed growth rates are connected by orange lines.

<https://doi.org/10.1371/journal.pcbi.1007066.g004>

mechanisms for this fundamental second-order tradeoff. The exact pathways involved can be determined in future experiments.

Alternative explanations of the rate-yield tradeoff have been proposed, including membrane [31, 32] and cytosolic crowding [33, 34]. It is difficult to rule out these alternative constraints on cell growth, and it may be that multiple constraints operate at the same time. However, it is encouraging to see that the ME-model can explain the complex relationship between  $\mu$ ,  $Y$ ,  $q_{ac}$ , and  $q_{glc}$  with only metabolic and proteome allocation constraints. In the future, it will be possible to extend ME-models with additional constraints. For example, it has been proposed that the unmodeled protein fraction (UPF) is growth-rate dependent, and thus existing proteome allocation models with fixed UPF are inaccurate [34]. If this is indeed the case, then SSME- and ME-models with cytosolic crowding constraints can be developed to fully represent the interplay between crowding, proteome allocation, and pathway selection.

## Discussion

The *E. coli* ME-model provides a mechanistic and predictive model of rate-yield tradeoffs. It successfully reconciles several experimental data sets: i) uptake titration at low growth [8], ii) batch culture at higher growth rates [5], and iii) ALE endpoint strains (this study). These data sets, when analyzed with the ME-model, show the existence of a two-dimensional rate-yield tradeoff. The first dimension (“d1”) rate-yield tradeoff is  $\mu$ -Y tradeoff and the second dimension (“d2”) is  $q_{glc}$ -Y tradeoff.

From a mathematical perspective, one can describe the observed tradeoffs as correlations between any pair of the four variables  $\mu$ ,  $Y$ ,  $q_{glc}$ , and  $q_{ac}$ . The two particular dimensions of the tradeoff that we describe,  $\mu$ -Y (“d1”) and  $q_{glc}$ -Y (“d2”), are motivated by two different trends in our physiological observations. First, the previously-reported strong linear correlation between  $\mu$  and  $Y$  [5] occurs for isogenic cultures under carbon limitation. The second dimension “d2” appears when comparing laboratory evolution endpoint strains, where  $q_{glc}$ ,  $q_{ac}$ , and  $Y$  are observed to vary at a fixed  $\mu$ , with a linear relationship in  $q_{glc}$ - $q_{ac}$  and a corresponding inverse proportional relationship in  $q_{glc}$ -Y. This two-dimensional tradeoff cannot be deciphered from simpler intuitive models, but it can be derived from the comprehensive set of metabolic and gene expression pathways represented by the ME-model.

Furthermore, this study employed a multi-scale modeling approach where a small-scale model was used to guide parameter estimation in the genome-scale ME-model. This

approach—which has been termed Tunable Resolution (TR) modeling [35]—was essential to the success of the study, and we expect that both small-scale and genome-scale models will continue to play an important role in understanding the genotype-phenotype relationship.

The two-dimensional rate-yield tradeoff appears as a result of ALE selection for  $\mu$  when alternative pathway selection strategies achieve the same growth rate. Proton leakage and alternative ETC pathway selection are plausible mechanisms for modifying the P/O ratio and creating the  $q_{glc}$ -Y tradeoff. In addition, the flux ratio between glycolysis (GAPD, *gapA*) and the pentose phosphate pathway (GND, *gnd*) might play a significant role in the  $q_{glc}$ -Y tradeoff. Those mechanisms can be tested experimentally. Finally, revealing the underlying regulation would be of great interest for establishing a deeper understanding of rate-yield tradeoffs. Combining ME-models with known regulatory mechanisms to explain cellular choices would achieve a long-standing goal in systems biology [36].

## Materials and methods

Phenotypic data including  $\mu$ ,  $q_{glc}$ ,  $q_{ac}$ , and excretion rates of other metabolic byproducts were collected for ALE endpoint strains (“1 Phenotypic characterization of *E. coli* strains” in [S1 Appendix](#)). In addition,  $^{13}\text{C}$  fluxes were measured from two of the strains with different growth rate and different glucose yield (“2  $^{13}\text{C}$  metabolic flux analysis” in [S1 Appendix](#)). Reference data points of rate-yield, growth-acetate relations of wild-type MG1655 and NCM3722 *E. coli* strains were collected from published studies [5, 8]. The coarse-grained proteome allocation model from [5] was reformulated as a small-scale ME-model (SSME-model, detail in “5 SSME-model parameters derivation” in [S1 Appendix](#)) and implemented by the COBRAME framework [16]. The genome-scale model iJL1678-ME was modified to fit experimental data by modifying the  $k_{eff}$  (enzyme turnover rate) of TCA cycle reactions, blocking target reactions, and modifying UPF (unmodeled protein fraction), GAM (growth associate maintenance energy), and NGAM (non-growth associate maintenance energy) (“8 Experimental data fitting” in [S1 Appendix](#)). Solution spaces were generated using flux balance analysis (incorporated in COBRAME) in the ME-model (“7 Solution space of the ME-model” in [S1 Appendix](#)). To determine the effect of modifying P/O ratio on ME-model solution spaces, a reaction representing proton leakage was added to the ME-model (“10 P/O ratio manipulation” in [S1 Appendix](#)). The effect of the *gnd* knockout was demonstrated by blocking the reaction GND in ME-model simulations (“11 *gnd* knockout simulation” in [S1 Appendix](#)).

## Supporting information

### **S1 Appendix. Detailed introduction and discussions of materials and methods.**

(PDF)

**S1 Table. Parameters comparison between two coarse-grained models.** Comparison between the coarse-grained proteome allocation model [5] and SSME-model. The derivation in detail is shown in “5 SSME-model parameter derivation” in [S1 Appendix](#).

(PDF)

**S2 Table. ME-model parameters.** Global parameter selection in iJL1678-ME model to fit the  $\mu$ -Y,  $\mu$ - $q_{ac}$  data as in [Fig 2C and 2D](#).

(PDF)

**S3 Table. iJL1678-ME model modification (blocked reactions).** Reactions that need to be turned off in the model to get quantitative fit of  $\mu$ -Y,  $\mu$ - $q_{ac}$  data as in [Fig 2C and 2D](#).

(XLSX)



**S4 Table. Essential exchanges.** Boundary reactions in the ME-model that need to be turned on.

(XLSX)

**S5 Table. Solution space variation (Below).** Reactions that after being turned off,  $q_{ac,min}$  increases. The variation of  $\mu$ - $q_{ac}$  solution space is shown in [S6B Fig](#).

(XLSX)

**S6 Table. Solution space variation (Above).** Reactions that after being turned off,  $q_{ac,max}$  decreases. The  $\mu$ - $q_{ac}$  solution space would varied as shown in [S6A Fig](#).

(XLSX)

**S7 Table. ALE phenotypes measurements.**  $\mu$ ,  $q_{ac}$  and  $q_{glc}$  measurements of the *E. coli* adapted MG1655 strains. Strains are replicates from [\[3\]](#).

(XLSX)

**S8 Table. *E. coli* K-12 MG1655 WT phenotypes measurements.**  $\mu$ ,  $q_{ac}$  and  $q_{glc}$  measurements. Data from [\[3\]](#).

(XLSX)

**S9 Table. *E. coli* NCM3722 glucose uptake titration phenotypes.** Data from [\[5\]](#).

(XLSX)

**S10 Table. *E. coli* K-12 MG1655 chemostat measurements.** Data from [\[8\]](#).

(XLSX)

**S11 Table.  $^{13}\text{C}$  metabolic flux analysis data.** Metabolic fluxes distribution of the highest  $\mu$  strain and highest  $Y$  strain among the ALE endpoint strains.

- Tab “Reactions”  $^{13}\text{C}$  MFA model and carbon mapping network.
- Tab “Net\_fluxes”  $^{13}\text{C}$  MFA calculated net fluxes. LB and UB are the 95% confidence intervals.
- Tab “SymMets”  $^{13}\text{C}$  MFA model symmetric metabolite carbon mappings.
- Tab “MS\_data” Measured mass distribution vectors (MDVs) by LC-MS/MS and their associated carbon mappings used for MFA calculations.
- Tab “Flux\_data” Measured uptake and secretion rates by HPLC.

(XLSX)

**S1 Fig. Scheme of the coarse-grained proteome allocation model [\[5\]](#).**

(TIF)

**S2 Fig. Modification of ME-model for fitting the experimental data, based on the guideline derived from SSME-model.** (A) Reduction of the enzyme efficiency for respiration ( $k_{eff,res}$ ) causes a more gradual acetate line. Reduction of UPF increases the model-predicted maximum  $\mu$ , shifting the acetate line to higher  $\mu$ . (B) Another approach of getting more gradual acetate line is to block bp1 reactions. (C) Activation of bp2 reactions (such as the Entner–Doudoroff pathway bypassing glycolysis) cause an inflection point and extension of the acetate line to higher  $\mu$ . (D) Workflow for the ME-model modification process. In the genome-scale ME-model, some TCA cycle reactions appeared as bp1 reactions, but, because they belong to the major respiration pathway of the cell, we will decreased their  $k_{effs}$  rather than blocking them entirely.

(TIF)

**S3 Fig. Summary of the modifications to the genome-scale ME-model.** (A) Compared to original *i*LL1678-ME, unmodeled protein fraction (UPF) is halved to 18%. (B) For the enzyme efficiency parameter  $k_{eff}$ , only the TCA  $k_{eff}$ s are modified. (C) The subsystems of the 24 bp1 reactions. (D) The subsystems of 26 bp2 reactions. (E) bp1 and bp2 reactions on the pathway map of central metabolism. (F) Acetate lines for the steps in the fitting process. More detailed illustration process is shown in [S4 Fig](#).

(TIF)

**S4 Fig. Iteration process of filling the ME-model prediction gap of growth rate dependent acetate excretion.** First two steps of bp1 iteration process are shown in the left two figures, where as we block the first bp1 reaction (ICL), the slope (threshold) of the acetate line drops. The changes of the threshold (bp1 modification) and acetate line end point (bp2 modification) from iteration Step 3–21 are shown in the right figure. Step 3–18 are the modification on bp1 reactions, where the threshold (red squares) gradually drops from high growth to low growth. Step 19–21 are the modification on bp2 reactions, where the acetate line end point (in red circles) drops. The blue solid line is the final prediction of  $\mu$ – $q_{ac}$  relation, which is the same as the blue line in [S3F Fig](#). More detail about bp1 and bp2 reactions are in [S3 Table](#).

(TIF)

**S5 Fig. By blocking byproduct excretion pathways in the ME-model, which is verified by the experimental data, the solution space was reduced from the pink region to the yellow region.**

(TIF)

**S6 Fig.  $\mu$ – $q_{ac}$  solution space variation in the ME-model.** Narrowing in the feasible range of alternative suboptimal solutions by blocking some target reactions. The new solution space after the variation is shown as the yellow in (A) and (B), with the original solution space in pink. (A) 24 target reactions ([S6 Table](#)) that are blocked where maximum  $q_{acs}$  in high  $\mu$  get lower, where the upper edge of the yellow region is below the upper edge of the pink region. The activation of one of these 24 reactions thus corresponding to higher  $q_{ac}$  with lower  $Y$ . (B) 11 target reactions ([S5 Table](#)) corresponding to lower  $q_{ac}$  with lower  $Y$ , blocking those reactions will get the minimum  $q_{ac}$  (lower edge of the yellow region) closed to the  $Y$ -maximized  $q_{ac}$  solution. (C) The method of picking reactions to block: Looking for the reactions that are not activated in the yield-maximized solution but activated at the maximal and minimal of the  $\mu$ – $q_{ac}$  solution space, where the principal is to keep the  $Y$ -maximized solutions unchanged.

(TIF)

**S7 Fig. Expansion of solution space from the SSME-model by adding model reactions.** The expanded part of the solution space is shown as yellow in (A)–(C), compared to the original SSME-model solution spaces are in blue. (A) All added reactions ((1)–(4) in D) expand the solution space to include lower- $Y$  solutions (B) Reactions (1) and (3) expand the solution space to low- $q_{ac}$  at high  $\mu$ . (C) Reactions (2) and (4) expand the solution space to high- $q_{ac}$  across all  $\mu$ . (D) Model reactions that are added in the SSME-model for expanding the original solution space, all those reactions are guaranteed not be activated in the  $Y$ -maximized solutions so that the  $Y$ -optimal solution remains the same to fit data from [5]. Reaction (1) corresponds to the reactions that would generate products other than acetate such as pyruvate excretion, lactate excretion, etc. Reaction (2) is representative to the reactions that would generate other products, but at the same time generating acetate, such as pyruvate formate lyase (PFL), which produce formate and acetyl-CoA (precursor of acetate) from pyruvate. Reaction (3) and (4) could both be referred from the futile cycle in energy production and consumption, where (3) are the reactions that are less efficient than the optimal pathway, such as the

alternative reactions in ETC which are less efficient in transporting electrons, while (4) are the reactions that would waste more energy in the same growth comparing to the optimal state, such as the reactions that would cause proton leakage.

(TIF)

## Acknowledgments

The authors would like to thank Ali Ebrahim, Laurence Yang, and Colton Lloyd for the assistance with ME-model analysis. Moreover, we are grateful to Ke Chen, David Heckmann, Marc Abrahms, Amitesh Anand and Brian Taylor for providing their advice on the manuscript. We highly appreciate Andrea De Martino for providing single strain (K-12 MG1655 and NCM3722) growth yield and acetate overflow data. And at last, we would like to express our extraordinary thanks to Terence Hwa and Matteo Mori for advice on developing multi-scale models of proteome allocation.

## Author Contributions

**Conceptualization:** Adam M. Feist, Bernhard O. Palsson, Zachary A. King.

**Data curation:** Chuankai Cheng, Edward J. O'Brien, Douglas McCloskey, Jose Utrilla, Connor Olson, Ryan A. LaCroix, Troy E. Sandberg, Zachary A. King.

**Formal analysis:** Chuankai Cheng, Edward J. O'Brien, Zachary A. King.

**Funding acquisition:** Adam M. Feist, Bernhard O. Palsson, Zachary A. King.

**Investigation:** Chuankai Cheng, Edward J. O'Brien, Douglas McCloskey, Ryan A. LaCroix, Troy E. Sandberg, Bernhard O. Palsson, Zachary A. King.

**Methodology:** Chuankai Cheng, Edward J. O'Brien, Bernhard O. Palsson, Zachary A. King.

**Project administration:** Bernhard O. Palsson, Zachary A. King.

**Resources:** Zachary A. King.

**Software:** Chuankai Cheng, Edward J. O'Brien, Zachary A. King.

**Supervision:** Bernhard O. Palsson, Zachary A. King.

**Validation:** Chuankai Cheng, Bernhard O. Palsson, Zachary A. King.

**Visualization:** Chuankai Cheng, Bernhard O. Palsson, Zachary A. King.

**Writing – original draft:** Chuankai Cheng, Bernhard O. Palsson, Zachary A. King.

**Writing – review & editing:** Chuankai Cheng, Bernhard O. Palsson, Zachary A. King.

## References

1. Lipson DA. The complex relationship between microbial growth rate and yield and its implications for ecosystem processes. *Front Microbiol.* 2015; 6:615. <https://doi.org/10.3389/fmicb.2015.00615> PMID: 26136742
2. Peebo K, Valgepea K, Nahku R, Riis G, Oun M, Adamberg K, et al. Coordinated activation of PTA-ACS and TCA cycles strongly reduces overflow metabolism of acetate in *Escherichia coli*. *Appl Microbiol Biotechnol.* 2014; 98(11):5131–5143. <https://doi.org/10.1007/s00253-014-5613-y> PMID: 24633370
3. LaCroix RA, Sandberg TE, O'Brien EJ, Utrilla J, Ebrahim A, Guzman GI, et al. Use of adaptive laboratory evolution to discover key mutations enabling rapid growth of *Escherichia coli* K-12 MG1655 on glucose minimal medium. *Appl Environ Microbiol.* 2015; 81(1):17–30. <https://doi.org/10.1128/AEM.02246-14> PMID: 25304508

4. Bekker M, de Vries S, Ter Beek A, Hellingwerf KJ, de Mattos MJT. Respiration of *Escherichia coli* can be fully uncoupled via the nonelectrogenic terminal cytochrome bd-II oxidase. *J Bacteriol.* 2009; 191(17):5510–5517. <https://doi.org/10.1128/JB.00562-09> PMID: 19542282
5. Basan M, Hui S, Okano H, Zhang Z, Shen Y, Williamson JR, et al. Overflow metabolism in *Escherichia coli* results from efficient proteome allocation. *Nature.* 2015; 528(7580):99–104. <https://doi.org/10.1038/nature15765> PMID: 26632588
6. Mori M, Hwa T, Martin OC, De Martino A, Marinari E. Constrained Allocation Flux Balance Analysis. *PLoS Comput Biol.* 2016; 12(6):e1004913. <https://doi.org/10.1371/journal.pcbi.1004913> PMID: 27355325
7. Molenaar D, van Berlo R, de Ridder D, Teusink B. Shifts in growth strategies reflect tradeoffs in cellular economics. *Mol Syst Biol.* 2009; 5:323. <https://doi.org/10.1038/msb.2009.82> PMID: 19888218
8. Nanchen A, Schicker A, Sauer U. Nonlinear dependency of intracellular fluxes on growth rate in miniaturized continuous cultures of *Escherichia coli*. *Appl Environ Microbiol.* 2006; 72(2):1164–1172. <https://doi.org/10.1128/AEM.72.2.1164-1172.2006> PMID: 16461663
9. O'Brien EJ, Lerman JA, Chang RL, Hyduke DR, Palsson BO. Genome-scale models of metabolism and gene expression extend and refine growth phenotype prediction. *Mol Syst Biol.* 2013; 9(1):693. <https://doi.org/10.1038/msb.2013.52> PMID: 24084808
10. Monod J. The Growth of Bacterial Cultures. *Annu Rev Microbiol.* 1949; 3(1):371–394. <https://doi.org/10.1146/annurev.mi.03.100149.002103>
11. Pirt SJ. The maintenance energy of bacteria in growing cultures. *Proc R Soc Lond B Biol Sci.* 1965; 163(991):224–231. <https://doi.org/10.1098/rspb.1965.0069> PMID: 4378482
12. Novak M, Pfeiffer T, Lenski RE, Sauer U, Bonhoeffer S. Experimental tests for an evolutionary trade-off between growth rate and yield in *E. coli*. *Am Nat.* 2006; 168(2):242–251. <https://doi.org/10.1086/506527> PMID: 16874633
13. Pfeiffer T, Schuster S, Bonhoeffer S. Cooperation and competition in the evolution of ATP-producing pathways. *Science.* 2001; 292(5516):504–507. <https://doi.org/10.1126/science.1058079> PMID: 11283355
14. Bachmann H, Fischlechner M, Rabbers I, Barfa N, Branco dos Santos F, Molenaar D, et al. Availability of public goods shapes the evolution of competing metabolic strategies. *Proc Natl Acad Sci U S A.* 2013; 110(35):14302–14307. <https://doi.org/10.1073/pnas.1308523110> PMID: 23940318
15. Sandberg TE, Pedersen M, LaCroix RA, Ebrahim A, Bonde M, Herrgard MJ, et al. Evolution of *Escherichia coli* to 42°C and subsequent genetic engineering reveals adaptive mechanisms and novel mutations. *Mol Biol Evol.* 2014; 31(10):2647–2662. <https://doi.org/10.1093/molbev/msu209> PMID: 25015645
16. Lloyd CJ, Ebrahim A, Yang L, King ZA, Catoiu E, O'Brien EJ, et al. COBRAme: A computational framework for genome-scale models of metabolism and gene expression. *PLoS Comput Biol.* 2018; 14(7): e1006302. <https://doi.org/10.1371/journal.pcbi.1006302> PMID: 29975681
17. Booth IR. Stress and the single cell: intrapopulation diversity is a mechanism to ensure survival upon exposure to stress. *Int J Food Microbiol.* 2002; 78(1-2):19–30. [https://doi.org/10.1016/S0168-1605\(02\)00239-8](https://doi.org/10.1016/S0168-1605(02)00239-8) PMID: 12222634
18. Elena SF, Lenski RE. Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nat Rev Genet.* 2003; 4(6):457–469. <https://doi.org/10.1038/nrg1088> PMID: 12776215
19. Barrick JE, Lenski RE. Genome dynamics during experimental evolution. *Nat Rev Genet.* 2013; 14(12):827–839. <https://doi.org/10.1038/nrg3564> PMID: 24166031
20. Utrilla J, O'Brien EJ, Chen K, McCloskey D, Cheung J, Wang H, et al. Global Rebalancing of Cellular Resources by Pleiotropic Point Mutations Illustrates a Multi-scale Mechanism of Adaptive Evolution. *Cell Syst.* 2016; 2(4):260–271. <https://doi.org/10.1016/j.cels.2016.04.003> PMID: 27135538
21. Ibarra RU, Edwards JS, Palsson BO. *Escherichia coli* K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. *Nature.* 2002; 420(November):20–23.
22. Varma A, Boesch BW, Palsson BO. Stoichiometric interpretation of *Escherichia coli* glucose catabolism under various oxygenation rates. *Appl Environ Microbiol.* 1993; 59(8):2465–2473. PMID: 8368835
23. Lee DH, Feist AM, Barrett CL, Palsson BØ. Cumulative number of cell divisions as a meaningful time-scale for adaptive laboratory evolution of *Escherichia coli*. *PLoS One.* 2011; 6(10):e26172. <https://doi.org/10.1371/journal.pone.0026172> PMID: 22028828
24. Lara AR, Caspeta L, Gosset G, Bolívar F, Ramírez OT. Utility of an *Escherichia coli* strain engineered in the substrate uptake system for improved culture performance at high glucose and cell concentrations: an alternative to fed-batch cultures. *Biotechnol Bioeng.* 2008; 99(4):893–901. <https://doi.org/10.1002/bit.21664> PMID: 17929322

25. Ebrahim A, Brunk E, Tan J, O'Brien EJ, Kim D, Szubin R, et al. Multi-omic data integration enables discovery of hidden biological regularities. *Nat Commun*. 2016; 7:13091. <https://doi.org/10.1038/ncomms13091> PMID: 27782110
26. Nilsson A, Nielsen J, Palsson BO. Metabolic Models of Protein Allocation Call for the Kinetome. *Cell Syst*. 2017; 5(6):538–541. <https://doi.org/10.1016/j.cels.2017.11.013> PMID: 29284126
27. Cheng C. The Study of Overflow Metabolism and kinetics in a model of metabolism and gene expression of *Escherichia coli*; 2017.
28. Russell JB, Cook GM. Energetics of bacterial growth: balance of anabolic and catabolic reactions. *Microbiol Rev*. 1995; 59(1):48–62. PMID: 7708012
29. Russell JB. The energy spilling reactions of bacteria and other organisms. *J Mol Microbiol Biotechnol*. 2007; 13(1-3):1–11. <https://doi.org/10.1159/000103591> PMID: 17693707
30. Jiao Z, Baba T, Mori H, Shimizu K. Analysis of metabolic and physiological responses to gnd knockout in *Escherichia coli* by using C-13 tracer experiment and enzyme activity measurement. *FEMS Microbiol Lett*. 2003; 220(2):295–301. [https://doi.org/10.1016/S0378-1097\(03\)00133-2](https://doi.org/10.1016/S0378-1097(03)00133-2) PMID: 12670695
31. Zhuang K, Vemuri GN, Mahadevan R. Economics of membrane occupancy and respiration-fermentation. *Mol Syst Biol*. 2011; 7(1):500. <https://doi.org/10.1038/msb.2011.34> PMID: 21694717
32. Szenk M, Dill KA, de Graff AMR. Why Do Fast-Growing Bacteria Enter Overflow Metabolism? Testing the Membrane Real Estate Hypothesis. *Cell Syst*. 2017; 5(2):95–104. <https://doi.org/10.1016/j.cels.2017.06.005> PMID: 28755958
33. Adadi R, Volkmer B, Milo R, Heinemann M, Shlomi T. Prediction of microbial growth rate versus biomass yield by a metabolic network with kinetic parameters. *PLoS Comput Biol*. 2012; 8(7):e1002575. <https://doi.org/10.1371/journal.pcbi.1002575> PMID: 22792053
34. Vazquez A, Oltvai ZN. Macromolecular crowding explains overflow metabolism in cells. *Sci Rep*. 2016; 6:31007. <https://doi.org/10.1038/srep31007> PMID: 27484619
35. Kirschner DE, Hunt CA, Marino S, Fallahi-Sichani M, Linderman JJ. Tuneable resolution as a systems biology approach for multi-scale, multi-compartment computational models: Tuneable resolution as a systems biology approach. *WIREs Syst Biol Med*. 2014; 6(4):289–309. <https://doi.org/10.1002/wsbm.1270>
36. Reed JL, Palsson B. Thirteen Years of Building Constraint-Based In Silico Models of *Escherichia coli*. *J Bacteriol*. 2003; 185(9):2692–2699. <https://doi.org/10.1128/JB.185.9.2692-2699.2003> PMID: 12700248